Navarro-Mateu Fernando (Orcid ID: 0000-0002-3228-623X)

The association of telomere length with substance use disorders: a systematic review and meta-analysis of observational studies.

Authors: Fernando Navarro-Mateu[1,2,3,4], Mathilde Husky[5], Pedro Cayuela-Fuentes[6], Francisco-Javier Álvarez[1], Agustín Roca-Vega[7], María Rubio-Aparicio[8], María Dolores Chirlaque[2,3,9,10], María Luisa Cayuela[11,12], Salvador Martínez[13,14] and Julio Sánchez-Meca[4]

*** Corresponding author:** Fernando Navarro-Mateu, MD, PhD. Unidad de Docencia, Investigación y Formación en Salud Mental (UDIF-SM). Servicio Murciano de Salud. c/ Lorca, nº 58. 30120-El Palmar (Murcia). Spain. Email: fernando.navarro@carm.es

Authors' information

[1] Unidad de Docencia, Investigación y Formación en Salud Mental (UDIF-SM). Servicio Murciano de Salud. Murcia. Spain.
[2] CIBER de Epidemiología y Salud Pública (CIBERESP). Madrid. Spain.
[3] IMIB-Arrixaca. Murcia. Spain.
[4] Departamento de Psicología Básica y Metodología. University of Murcia., Murcia. Spain.
[5] Université de Bordeaux, Laboratoire de Psychologie EA4139, Bordeaux, France.
[6] Escuela Universitaria de Enfermería de Cartagena, University of Murcia, Murcia. Spain.
[7] Biblioteca Virtual MurciaSalud. Centro Tecnológico de Información y Documentación Sanitaria, Servicio Murciano de Salud. Murcia. Spain.
[8] Departamento Psicología de la Salud, University of Alicante, Alicante. Spain.
[9] Servicio de Epidemiología. Consejería de Salud. Murcia. Spain.
[10] Departamento de Ciencias Sociosanitarias. University of Murcia. Spain.
[11] Grupo Telomerasa, Cáncer y Envejecimiento. Hospital Clínico Universitario Virgen de la Arrixaca. Murcia. Spain.
[12] CIBER de Enfermedades Raras (CIBERER). Madrid. Spain.
[13] Instituto de Neurociencias. UMH-CSIC. Alicante. Spain.
[14] CIBER de Salud Mental (CIBERSAM). Madrid. Spain.

Email addresses:

FNM: Fernando Navarro-Mateu. Email address: fernando.navarro@carm.es
MH: Mathilde Husky. Email address: mathilde.husky@u-bordeaux.fr
PCF: Pedro Cayuela-Fuentes. Email address: cayuela.pedro@gmail.com
FJA: Francisco-Javier Álvarez. Email address: franciscoj.alvarez@carm.es
ARV: Agustín Roca-Vega. Email address: agustin.roca@carm.es
MRA: María Rubio-Aparicio. Email address: maria.rubio@ua.es
MDC: María Dolores Chirlaque. Email address: mdolores.chirlaque@carm.es

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/add.15312

MLC: María Luisa Cayuela. Email address: marial.cayuela@carm.es
SM: Salvador Martínez. Email address: smartinez@umh.es
JSM: Julio Sánchez-Meca. Email address: jsmeca@um.es

Running head: Telomere length and substance use disorders.
Word count: 3938.
Conflict of interest declaration Dr Navarro-Mateu reports non-financial support from Otsuka outside the submitted work. Others coauthors declare no conflict of interest.

ABSTRACT

Background and Aims Several recent studies have investigated the relationship between telomere length and substance use disorders with inconsistent results. We aimed to assess this association and to identify moderators of the relationship.

Methods Systematic review and meta-analysis. Selection criteria were observational studies reporting telomere length in persons with a substance use disorder compared with a control group. Studies focused solely on nicotine addiction, employing other study designs, and non-human studies were excluded. Study selection and data extraction were independently conducted by two researchers following a standardized protocol and included studies up until December 2019. Standardized mean differences were used as the effect size index (*d*; 95%CI) and random-effects models were used for the meta-analysis. Cochran's *Q*-statistic, I^2 index, visual inspection of the forest plot, and a 95% prediction interval were applied to verify study heterogeneity. Subgroup analyses and meta-regressions were conducted to explore heterogeneity. Small study effects were examined using the "funnel plot", the Egger test, Duval and Tweedie's trim-and-fill method, and the PET-PEESE method. The risk of bias and the quality of evidence were assessed.

Results Ten studies (12 analysis units with 2,671 cases and 4,532 controls) met the selection criteria. An overall effect size of moderate magnitude was found (d_+= -0.63; 95%CI: -1.00 and -0.26; p=.0008). A potential small study effect was detected, as well as large heterogeneity between studies (*Q*-statistic p<.001, I^2=97.3%). Selection of controls, reporting laboratory

quality control procedures and total sample size significantly affected the effect size. The quality of the evidence was very low, based on risk of bias analysis and the GRADE system.

Conclusions People with substance use disorders appear to have shorter telomere length than controls; however, this finding should be interpreted with caution due to the poor quality of the evidence.

Systematic review registration: PROSPERO registration number CRD42019119785.

KEYWORDS Telomere length; Substance Use Disorders; Cellular aging; Alcohol; Meta-analysis; Systematic Review

INTRODUCTION

Telomeres are repetitive noncoding DNA protein structures consisting of nucleotide sequences of tandem TTAGGG repeats at the end of chromosomes in association with a protein complex. These structures are essential for maintaining genome stability (1) and for ensuring the regulation of gene expression (2). Telomere length (TL) varies throughout the lifespan and is considered as a marker of cellular aging (3–5). Telomere attrition has been associated with increased all-cause mortality risk (6) and in particular, with increased morbidity of various age-related diseases (7–12). Results of recent meta-analyses suggest that TL might be associated with a variety of mental disorders (13–19). However, a non-systematic review has highlighted inconsistencies of the published results regarding the association between substance use disorders (SUDs) and telomere length (20).

SUDs constitute one of the major public health issues around the world (21,22), are major contributors to burden of disease (23) with greater risk of disability (24) and mortality (25). Early detection of addiction is considered crucial for preventing premature morbidity and mortality (26). Comorbidity is highly prevalent between SUDs and both psychiatric disorders (27) and medical conditions (28). To the best of our knowledge, no systematic review or meta-analysis examining the association of telomere length with SUD related to any substance other than tobacco (29) was ever conducted.

The aims of the present study were i) to determine whether persons with SUDs have shorter telomere lengths as compared to healthy controls, ii) to explore potential differential effects with regard to diverse substances, iii) to identify potential moderators of the telomere length effect. The research questions were: i) Do people with SUDs have shorter telomere lengths compared with healthy controls?; ii) Are there differences in the association of TL with SUD as a function of the type of substance that is misused?; and iii) If heterogeneity is confirmed, what are the factors implicated?.

METHODS

Protocol and registration

The protocol of this investigation was registered with the International Prospective Register of Systematic Reviews (PROSPERO 2019 CRD42019119785, https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=119785) and published previously (30). We wrote this report using the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA guidelines) (31) and the proposal for reporting Meta-analysis of Observational Studies in Epidemiology (MOOSE) (32).

Study eligibility criteria

Inclusion criteria were as follows: (a) *Populations:* Adult with SUDs, except if the disorder is exclusively based on nicotine addiction, and healthy controls; (b) *Exposure*: SUD covered alcohol, illicit drugs including cocaine, opiates, or other substances (e.g. marijuana and amphetamine, among others). Case status had to be defined as having any substance use disorder identified through a clinical interview or using established standard diagnostic instruments including, but not limited to, the Structured Clinical Interview for DSM-IV (SCID), Computerized National Institute of Mental Health Diagnostic Interview Schedule (CDISIV), the Composite International Diagnostic Interview (CIDI) or any other diagnostic instrument based on Diagnostic and Statistical Manual of Mental Disorders (DSM-IV) criteria, or; (c) *Control group*: adults with no SUD diagnosis (e.g. from the general population, the community, unexposed out-patient, or hospital-based controls); (d) *Outcomes*: telomere length measurements with a detailed description of both the methods of measurement and the isolated tissue that was used; and (e) *Study designs*: observational studies (case-control, cohort, cross-sectional, longitudinal designs). Exclusion criteria were: systematic or narrative reviews, meta-analyses, studies with non-human samples or other designs including reviews, case-only

studies, family-based designs, population-based studies with healthy subjects only, as well as studies focused on tobacco smoking.

Information sources and search strategy

Comprehensive electronic searches were conducted to identify studies indexed in PubMed/MEDLINE, EMBASE, Psychlit/PsychINFO, and Web of Science databases (Web of Knowledge) from inception until December 2019. The search was performed by a librarian with expertise in systematic reviews. The following search terms were used for SUDs: "drug, substance, addiction, alcohol*, heroin, cocaine, opium, opioid, methamphetamine, morphine" and for telomeres: *Telomeres, telomerase, and telo**" (see Table S1). The references cited in each study included in this initial selection and in review articles were then manually searched to identify other potentially eligible studies. To minimize potential publication bias, both published and unpublished papers were eligible for inclusion. In order to identify unpublished studies, e-mails were sent to the corresponding authors of the selected studies to enquire about any potential study that met eligibility criteria. In the search strategy, no restrictions were placed on time period, sample size, ethnicity or language of publication.

Data extraction

The following data were extracted from each study following the previously defined protocol: i) identification data of the study (author(s), journal, language, and year of publication); ii) methods (study design, sample sizes for both cases and controls, diagnostic tools for the determination of case status, definition of case status, variables adjusted for in the analyses, attrition for cases and controls and differential attrition); iii) risk of bias assessment (described in greater detail below); iv) sample characteristics for both cases and controls separately (gender ratio, mean age and standard deviation (SD), ethnic background, education level, type of substance used in SUD cases, duration of SUD in cases, presence of comorbid mental disorders or medical conditions in cases, smoking status, exposure to childhood

adversities and other stressful life events); v) telomere-related information (telomere length, tissue source and telomere measurement method), and (vi) extrinsic characteristics (relevant ethical approval, conflict of interest disclosure and funding source).

If an article reported two or more studies with independent samples, then each independent study was included as an analysis unit in the meta-analysis. When essential data were unavailable in the original studies, authors of the respective papers were contacted and asked to provide additional data. Two reviewers independently determined eligibility and extracted data from included studies. Disagreements were resolved by consensus or reached with the involvement of a third reviewer. To assess the reliability of the data extraction process, in terms of inter-rater agreement, kappa coefficients were calculated between the two reviewers.

Risk of bias assessment

The risk of bias of each included study was assessed using the Newcastle-Ottawa Scale (NOS; (33)). Discrepancies in the quality assessment of each study were resolved by consensus. A Total Quality Score (TQS) of each individual study was calculated by adding all the stars (range: 0-9, with a higher score indicating higher overall quality). Studies were not weighted by the TQS and the influence on the effect size of each item was individually assessed (34). In addition to the NOS, several study characteristics (e.g. if a blind assay assessment and genetic quality procedures were reported, as well as the evaluation of psychiatric or physical comorbidities or the exposure to childhood adversities or other stressful events) were extracted to analyze their potential risk of bias on the effect sizes.

Effect size index

For each study, means and SDs on TL measured in T/S ratio scale were extracted. These data were converted into Hedges' standardized mean difference (*d*) as effect size index. The *d* index was calculated as the mean difference in telomere length between the SUD and control

groups, divided by the pooled standard deviation of the two groups (35). Negative *d*s represented a shorter telomere length for the SUD group as compared to the control group. By convention, *d* indices of 0.20, 0.50, and 0.80 (in absolute value) were considered of small, moderate, and large magnitude, respectively (36). For each *d* index, a 95% confidence interval (95% CI) was calculated. In this meta-analysis unadjusted effect sizes (*d*s) were used. As described in the Results section, the reason for not analyzing adjusted effect sizes was that the majority of the studies did not report the statistical information needed to calculate an adjusted effect size using the same metric used for the unadjusted standardized mean difference (i.e., adjusted means and SDs to calculate adjusted standardized mean differences). The potential influence of confounding factors was assessed as described below. Table S2 describes how the data were extracted from the studies and *d* indices were calculated.

Statistical analyses

Random-effects models were used to analyze the TL-SUD association due to an expectation of a high level of heterogeneity among the studies. An average effect size and a 95% CI was calculated with the improved method proposed by Hartung and Knapp (37–39). In addition, a 95% prediction interval around the average effect size was calculated in order to provide a prediction of the expected true effects if a new study is conducted (40).

To estimate heterogeneity between studies, the Cochran's *Q*-statistic, the I^2 index, and visual inspection of the forest plots were used. In addition, heterogeneity was assessed with the between-studies variance and corresponding 95% confidence interval (41). Finally, the estimated proportion (and 95% confidence interval) of true effect sizes exceeding a meaningful threshold was calculated, considering -0.20 as the threshold effect size for these calculations in terms of standardized mean difference (42).

In cases of moderate-to-large heterogeneity (I^2> 25%), we attempted to identify possible explanations using subgroup analyses and meta-regressions based on the most

important characteristics of the studies, including items used to evaluate the risk of bias. The analyses of moderating variables was individually assessed (32) and were accomplished by assuming a mixed-effects model (43). The improved *F* statistic was applied for testing the statistical significance of each moderator (44). To estimate the proportion of variance accounted for by the moderator, an R^2 index was calculated (45). Simple and multiple mixed-effects meta-regression was applied to analyze the influence of the following moderators on the effect sizes: publication year, mean and SD of the age (total, case, and control samples), mean age difference, SD of age difference, percent male (total, case, and control sample), percent male difference, sample size, and NOS total quality score.

The presence of small study effects was examined using the "funnel plot" method in combination with Duval and Tweedie's trim-and-fill method (46), the Egger test (47), and the Precision-Effect Test–Precision-Effect Estimate with Standard Error (PET-PEESE) method (48). An additional sensitivity analysis was performed with the 'leave-one-out' method, by systematically removing each study and recalculating the overall results. All statistical analyses were conducted using the metafor program in R (49), except for the PET-PEESE method that was conducted with SPSS macros (48). The Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach was used to evaluate the quality of evidence (50).

RESULTS

Study eligibility and data collection

We first identified a total of 1,173 studies. After duplicates were removed, titles and abstracts of 701 studies were screened for eligibility and 558 were excluded. A total of 143 full-text studies were assessed for eligibility and 133 of them were excluded (see flow chart in Figure 1 and individual reasons for exclusion in Table S3). Inter-rater agreement in the selection process was reached in 96% of the studies. Finally, 10 studies (12 analysis units) were

selected for the meta-analysis. Although efforts to identify unpublished studies were made, all of the studies included in this meta-analysis were published articles. The main characteristics of these studies are summarized in Tables 1 and 2. The median (SD) of the Cohen's kappa inter-rater agreement coefficient was 0.70 (0.24) and ranged from 0.16 to 1.00.

The 10 eligible studies included 7,203 participants (2,671 cases and 4,532 controls). As shown in Table 1, all studies applied a case-control design. The most represented countries were the U.S. with 3 studies (51–53) and Japan with 2 studies (54,55). Case samples presented mean ages ranging between 26.2 and 74.5 years old (Mean= 47.4), whereas control samples ranged from 33.3 to 75.1 (Mean= 55.1). Three studies included men only (55–57) and one study with two analysis units included women only (58). Related to the type of substance used, five studies investigated alcohol (52–55,57), one alcohol and cocaine (51), one cocaine (58), one tobacco and marijuana (56), one a mixture of cocaine, heroin, methamphetamine and morphine (59), and one study did not describe the substances consumed by those diagnosed with SUD (60).

Adjusted effect sizes were not calculated due to the majority of the studies did not report the statistical data needed to obtain adjusted standardized mean differences (51-53,58-60). In addition, in one study this information was reported (57), but in terms of geometric means and not as arithmetic means, and two studies did not apply adjusted analyses (54,56). Only one study (55) reported statistical data needed to calculate an adjusted standardized mean difference, with a value of d_{adj} = -1.81 (95%CI = -2.10 and -1.52), which was very similar to the unadjusted *d* index: *d* = -1.89 (95%CI = -2.18 and -1.59). As shown in Table 3, the variables most frequently used to adjust the SUD-TL association were the age, sex, and smoking status. The results of the adjusted statistical analyses reported in the studies (multiple linear regression models in most cases) are described in Table S3. Therefore, meta-analytic calculations were based on unadjusted effect sizes.

Average effect size and heterogeneity

A forest plot of the *d* indices comparing average telomere length of SUD and control samples is presented in Figure 2. With one exception (60), every study exhibited shortened telomere length in SUD samples in comparison with controls, with eight studies reaching statistical significance (52,54–60). A overall effect size of moderate magnitude was found (d_+ = -0.63; 95%CI: -1.00 and -0.26; p = .0008). The 95% prediction interval (-2.06 to 0.80) was wide, indicating that the expected effect size in a new study could exhibit a wide range of true effect sizes, both of negative or positive sign. It was estimated, taking into account the overall effect size and the between-studies variance, that about 75.5% (95%CI: 54.6%, 96.4%) of true effect sizes exceeded the threshold for a scientifically meaningful size of d = -0.20. In addition, taking d = 0.20 as a threshold in the inverse direction, this method estimated that only 9.2% of true effect sizes exceeded that threshold (95%CI: 0%, 23%). As a sensitivity analysis, the 'leave-one-out' method was applied, finding three studies whose exclusion led to a change larger than 10% in the overall effect size (d_{-1} values = -0.56 (56); -0.50 (55)), and -0.72 (60)), but in all cases the adjusted overall effect size was statistically significant and of moderate magnitude (d > |0.50|).

The *Q*-statistic to assess heterogeneity among the effect sizes was statistically significant [$Q(11)$ = 256.56, p<.001], and the I^2 index was of large magnitude (I^2 = 97.3%), as well as the between-studies variance (τ^2 = 0.39; 95%CI: 0.04, 0.74). Taken together, these findings revealed the existence of large heterogeneity between studies.

Small study effect analyses

To assess whether small study effects were affecting to the meta-analytic results, a funnel plot was constructed as reported in Figure 3. The existence of asymmetry in the funnel plot was corroborated with the Egger test, that reached statistical significance [$t(10)$ = -1.93, p=.082]. The trim-and-fill method to symmetrize the funnel plot did not add to the effect size.

However, when the PET-PEESE method was applied, an estimate of the overall effect size adjusted by small study effects was of practically null magnitude (d_{PET} = 0.05; 95%CI: -0.46, 0.56).

Risk of bias analyses

The methodological quality of the studies was assessed with the Newcastle-Ottawa Scale (NOS), together with several additional items not included in NOS (see Table 3 and Figure 4). According to the GRADE system (50), there is very low quality evidence that people with SUDs have shorter TL (see Table S4, based on Cochrane's template for assessing the GRADE criteria (61)).

The potential relationship between each item of NOS and the effect sizes was assessed by means of subgroup analyses (see Table 4). There was some evidence for the effect size varying by the selection of controls (p = .016; R^2 = .43). Studies that selected controls from a hospitalized population or with no description of the selection process exhibited a slightly higher but non-statistically significant average effect size as compared to those with community controls (d_+ = -0.94 vs. -0.07). Table 4 presents the results of subgroup analyses for three additional methodological characteristics. Of these analyses, the only one that exhibited a relevant association with the effect sizes was whether the study reported quality control procedures in genotyping methods (p = .028; R^2 = .37), such that a lower average effect size was found when quality control methods were applied than when they were not reported (d_+ = -0.50 vs. -1.88). However, this result must be interpreted cautiously because only one study did not report quality control methods.

Types of substances related to SUD and telomere length

Studies were classified in three categories as a function of the type of substance misuse: alcohol, other substances (mainly cocaine), and alcohol plus cocaine. Table 5 presents the results of comparing the average effect sizes for these three categories. No relevant differences

were found between the three types of substance (p = .788; R^2 = 0). An additional analysis consisted of defining three dichotomous variables to categorize studies included consumers of alcohol, cocaine, and other substances, with codes 0 (No consumers of that substance) and 1 (consumers). Then, a multiple meta-regression analysis was applied with these three moderators and the effect sizes as the dependent variable. There was no evidence of a relationship between the type of substance and the effect sizes (F(3, 7) = 0.47, p = .715, R^2 = 0).

Study techniques of telomere length measurement and SUD determination

SUD status was assessed by clinical interview or through self-reported instruments. As shown in Table 5, no relevant differences were found between the two methods of SUD assessment (p = .280, R^2 = .02), although the magnitude of the difference in TL between SUD and controls was larger when cases were assessed by clinical interview (d_+ = -0.73 vs -0.18). A smaller effect size was found when TL was measured using qPCR (quantitative polymerase chain reaction) method (d_+ = -0.55 vs. -0.87), although not reaching statistical significance (p = .482, R^2 = 0). Differences in source tissue used in the biological samples to measure TL did not exhibit a relevant association with the effect sizes (p = .953, R^2 = 0).

Analysis of additional moderating variables

Subgroup analyses (ANOVA) were conducted to investigate the potential relationships between clinical, socio-demographical, and contextual characteristics and the effect sizes (Table 5). Neither the presence of psychiatric comorbidity (p = .415, R^2 = 0), medical comorbidity (p = .660, R^2 = 0), childhood trauma (p = .771, R^2 = 0), or exposure to other stressful events (p = .917, R^2 = 0) exhibited a relevant relationship with the effect sizes. Ethnicity of the sample (p = .788, R^2 = 0), country of residence (p = .114, R^2 = .61) or continent (p = .357, R^2 = .09) where the study was conducted, or funding type (p = .196, R^2 = .08) did not seem to affect the TL-SUD association either.

Meta-regressions were applied to assess the influence of unbalanced distribution of several socio-demographic moderators on the TL-SUD association. As shown in Table 6, none of them reached a relevant association with the effect sizes: mean age and SD of the samples (total, cases, and controls), percent males, or study publication year, all of them exhibiting percentages of variance accounted for lower than 10%. However, the total sample size of the studies exhibited a strong relationship with the effect sizes, with 54% of variance accounted for (p = .007; R^2 = .54; see Table 6). Figure S1 presents a scatter plot of how sample size affected the TL-SUD association. In particular, studies with small sample sizes exhibited stronger TL-SUD associations than studies with larger sample sizes. In other words, studies with small sample sizes found that SUD samples presented shortened TL in a larger magnitude than studies with large sample sizes. This result was coherent with the result of the Egger test above described.

DISCUSSION

To the best of our knowledge, this systematic review and meta-analysis is the first to systematically assess the TL-SUD association. The main result of a total of 12 analysis units suggests that people diagnosed with a SUD have a shorter TL as compared to controls. This finding is consistent with other recent meta-analyses suggesting that a shorter TL is associated with i) other mental disorders (19), such as depression (13,14), post-traumatic stress disorder (17), anxiety (62), and schizophrenia (15,18); ii) cigarette smoking (29), and ii) with other chronic age-related diseases, such as metabolic syndrome (7), diabetes mellitus (8), hypertension (9), cardiometabolic outcomes (10) and cardiovascular disease (11), and Alzheimer's disease (12).

Several strengths of our study should be highlighted. First, data on several potential moderating factors (e.g., childhood adversities, exposure to other stressful events, and psychiatric and physical comorbidities) was evaluated. Second, quality assessment was

implemented (63) using the Ottawa-Newcastle scale (33) and, although a TQS was calculated, each item was assessed individually in their influence on the magnitude of the effect (32). Third, we have evaluated risk of bias (63) and applied GRADE criteria to assess the quality of evidence (50). Finally, we have used PRISMA (31) and MOOSE checklist when writing this report (32), the protocol was registered in PROSPERO and has recently been published (30).

Nevertheless, some limitations deserve careful consideration. At the study level these were: firstly, some difficulties to extract some characteristics from the studies due to incomplete reporting and a very low quality of evidence based on GRADE criteria (50). The Strengthening the REporting of Genetic Association studies (STREGA) Statement was published in 2009 (64) as an extension of the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) (65) and was specifically designed to enhance the transparency of the reports of genetic association studies based on observational designs. While all ten studies were published afterwards, none of them has followed these international recommendations. Secondly, the scarce number of included studies limited the ability to identify potential moderators of the association. In our attempts to explain the large heterogeneity observed, only two methodological characteristics were identified as moderators of the TL-SUD association. However, other factors have been previously described (e.g. childhood adversities (66), exposure to other stressful events (67), cigarette smoking (29), physical (7–12) and psychiatric comorbidities (13–19)). Moreover, concerns about the impact of different measurement techniques and variability in several critical methodological steps in measuring TL which may vary between cases and controls, such as sample type selection, protocol of sample collection, storage, processing issues, the lapse of time between sample collection and analyses, and assay procedures among others, have been recently published (68–70). As a consequence, in an effort to improve the quality of telomere length research, a checklist of the minimum critical information necessary to enhance reproducibility between

laboratories, reliability and methodological rigor has been proposed (70). Thirdly, small study effect is suggested by our analyses, such that the most precise studies (i.e., with large sample sizes) were those that exhibited a very weak TL-SUD association, whereas studies with small sample sizes were those that obtained the largest TL-SUD associations. And fourthly, all were case-control studies except two studies (with three analysis units) that were cohorts in design but used a nested case-control analysis (53,71) with TL measured at a single point in time. Only one of the latter, The Heart and Soul Study described in (53), measured TL in a prospective manner, although the median absolute change in TL was not significant between alcohol consumers and abstainers after 5-year follow-up.

At the review level, the analyses were based on unadjusted effect estimates. Using unadjusted effect estimates in place of adjusted estimates can lead to biased estimates of meta-analytic parameters, such that the results must interpreted with caution. Another limitation was that the scarcity of studies limited subgroup or stratified analyses of individual substances. In addition, the results of the analyses must be interpreted with caution due to the large number of moderating variables analyzed and the small number of studies meta-analyzed.

Finally, the causal nature of the association between SUDs and TL needs to be interpreted with caution due to other potential explanations and limitations of current research on this topic. A plausible mechanism is that consumption of illicit drugs might misbalance the equilibrium of telomere addition by telomerase, and telomere attrition due to DNA end replication and other factors, e.g. stressful experiences elevating oxidative stress (72,73). However, this traditional causal explanation of the association of a shorter TL and SUDs has recently been questioned (74). Telomeres are specialized structures and their complex functionality still needs to be well-understood, as they cannot be considered as a passive marker of aging, but also as essential for genome stability and its protection as well as implicated in its expression (1,2).

Future research should improve several aspects in designing and reporting studies (e.g. state in the method section that the TL measurements were assessed blind to the condition of participants and to warrant that controls pertain to the same population than cases). Longitudinal studies are needed to establish a temporal relationship between TL and SUDs and to contribute to the clarification of the nature and direction of the relationship. High-quality prospective studies with larger samples will contribute to ascertain the complex nature of the relationship between shortened TL in SUDs. Finally, relevant statistical information is very frequently missing in the studies; in particular, adjusted means and SDs. Studies should report adjusted effect estimates to improve the interpretability of their results.

In summary, we have demonstrated that a shortened TL is associated to SUDs. Though noteworthy, caution should be kept in mind when interpreting these results as several methodological issues may alternatively explain these findings. If confirmed, TL is a promising marker of accelerated biological ageing in people with SUDs, a potential biomarker for prevention of premature morbidity and mortality and as a viable predictor of different pharmacological (75–77) and non-pharmacological (78,79) interventions.

DECLARATIONS SECTION

Ethical Approval and Consent to participate Not applicable.

Consent for publication Not applicable.

Data access statement: All data generated and datasets used during the current study will be available from the corresponding author on reasonable request.

Funding The publication of this study is supported by the "Observatorio sobre Drogas de la Región de Murcia" and IMIB-Arrixaca as part of the PEGASUS-Murcia (Psychiatric Enquiry to General Population in Southeast Spain-Murcia) project. However, the institutions have not been involved in the design, field work, analyses or the writing of the manuscript.

Authors' contributions FNM is the guarantor. All authors contributed to the development of the selection criteria, the risk of bias assessment strategy and to the design of the protocol. ARV and FNM developed the search strategy. PCF and FJA extracted data. MDC participated in discrepancies resolution. MLC and SM provided expertise on telomere length. MRA, FNM and JSM performed statistical analyses. FNM, JSM and MH wrote the draft and all authors have read and approved the final version of the manuscript.

Acknowledgments Not applicable

REFERENCES

1. Blackburn EH, Epel ES, Lin J. Human telomere biology: A contributory and interactive factor in aging, disease risks, and protection. Science. 2015;350(6265):1193–8.

2. Blackburn EH. Telomeres and telomerase: their mechanisms of action and the effects of altering their functions. FEBS Lett. 2005 Feb 7;579(4):859–62.

3. Aubert G, Lansdorp PM. Telomeres and aging. Physiol Rev. 2008 Apr;88(2):557–79.

4. Bernadotte A, Mikhelson VM, Spivak IM. Markers of cellular senescence. Telomere shortening as a marker of cellular senescence. Aging. 2016 Jan;8(1):3–11.

5. Zhu Y, Liu X, Ding X, Wang F, Geng X. Telomere and its role in the aging pathways: telomere shortening, cell senescence and mitochondria dysfunction. Biogerontology. 2019;20(1):1–16.

6. Wang Q, Zhan Y, Pedersen NL, Fang F, Hägg S. Telomere Length and All-Cause Mortality: A Meta-analysis. Ageing Res Rev. 2018 Dec;48:11–20.

7. Cheng Y-Y, Kao T-W, Chang Y-W, Wu C-J, Peng T-C, Wu L-W, et al. Examining the gender difference in the association between metabolic syndrome and the mean leukocyte telomere length. PloS One. 2017;12(7):e0180687.

8. Wang J, Dong X, Cao L, Sun Y, Qiu Y, Zhang Y, et al. Association between telomere length and diabetes mellitus: A meta-analysis. J Int Med Res. 2016 Dec;44(6):1156–73.

9. Tellechea ML, Pirola CJ. The impact of hypertension on leukocyte telomere length: a systematic review and meta-analysis of human studies. J Hum Hypertens. 2017 Feb;31(2):99–105.

10. D'Mello MJJ, Ross SA, Briel M, Anand SS, Gerstein H, Paré G. Association between shortened leukocyte telomere length and cardiometabolic outcomes: systematic review and meta-analysis. Circ Cardiovasc Genet. 2015 Feb;8(1):82–90.

11. Haycock PC, Heydon EE, Kaptoge S, Butterworth AS, Thompson A, Willeit P. Leucocyte telomere length and risk of cardiovascular disease: systematic review and meta-analysis. BMJ. 2014 Jul 8;349:g4227.

12. Forero DA, González-Giraldo Y, López-Quintero C, Castro-Vega LJ, Barreto GE, Perry G. Meta-analysis of Telomere Length in Alzheimer's Disease. J Gerontol A Biol Sci Med Sci. 2016;71(8):1069–73.

13. Ridout KK, Ridout SJ, Price LH, Sen S, Tyrka AR. Depression and telomere length: A meta-analysis. J Affect Disord. 2016 Feb;191:237–47.

14. Schutte NS, Malouff JM. The association between depression and leukocyte telomere length: a meta-analysis. Depress Anxiety. 2015 Apr;32(4):229–38.

15. Polho GB, De-Paula VJ, Cardillo G, dos SB, Kerr DS. Leukocyte telomere length in patients with schizophrenia: A meta-analysis. SchizophrRes. 2015 Jul;165(2–3):195–200.

16. Lin P-Y, Huang Y-C, Hung C-F. Shortened telomere length in patients with depression: A meta-analytic study. J Psychiatr Res. 2016 May;76:84–93.

17. Li X, Wang J, Zhou J, Huang P, Li J. The association between post-traumatic stress disorder and shorter telomere length: A systematic review and meta-analysis. J Affect Disord. 2017 15;218:322–6.

18. Rao S, Kota LN, Li Z, Yao Y, Tang J, Mao C, et al. Accelerated leukocyte telomere erosion in schizophrenia: Evidence from the present study and a meta-analysis. J Psychiatr Res. 2016;79:50–6.

19. Darrow SM, Verhoeven JE, Révész D, Lindqvist D, Penninx BWJH, Delucchi KL, et al. The Association Between Psychiatric Disorders and Telomere Length: A Meta-Analysis Involving 14,827 Persons. Psychosom Med. 2016;78(7):776–87.

20. Monroy-Jaramillo N, Dyukova E, Walss-Bass C. Telomere length in psychiatric disorders: Is it more than an ageing marker? World J Biol Psychiatry. 2017;1–19.

21. Degenhardt L, Whiteford HA, Ferrari AJ, Baxter AJ, Charlson FJ, Hall WD, et al. Global burden of disease attributable to illicit drug use and dependence: findings from the Global Burden of Disease Study 2010. Lancet Lond Engl. 2013 Nov 9;382(9904):1564–74.

22. Whiteford HA, Ferrari AJ, Degenhardt L, Feigin V, Vos T. The Global Burden of Mental, Neurological and Substance Use Disorders: An Analysis from the Global Burden of Disease Study 2010. PLoS One. 2015 Feb 6;10(2):e0116820.

23. Laramée P, Kusel J, Leonard S, Aubin H-J, François C, Daeppen J-B. The economic burden of alcohol dependence in Europe. Alcohol Alcohol Oxf Oxfs. 2013 Jun;48(3):259–69.

24. Samokhvalov AV, Popova S, Room R, Ramonas M, Rehm J. Disability associated with alcohol abuse and dependence. Alcohol Clin Exp Res. 2010 Nov;34(11):1871–8.

25. Laramée P, Leonard S, Buchanan-Hughes A, Warnakula S, Daeppen J-B, Rehm J. Risk of All-Cause Mortality in Alcohol-Dependent Individuals: A Systematic Literature Review and Meta-Analysis. EBioMedicine. 2015 Oct;2(10):1394–404.

26. Bachi K, Sierra S, Volkow ND, Goldstein RZ, Alia-Klein N. Is biological aging accelerated in drug addiction? Curr Opin Behav Sci. 2017 Feb;13:34–9.

27. Lai HMX, Cleary M, Sitharthan T, Hunt GE. Prevalence of comorbid substance use, anxiety and mood disorders in epidemiological surveys, 1990-2014: A systematic review and meta-analysis. Drug Alcohol Depend. 2015 Sep 1;154:1–13.

28. Onyeka IN, Collier Høegh M, Nåheim Eien EM, Nwaru BI, Melle I. Comorbidity of Physical Disorders Among Patients With Severe Mental Illness With and Without Substance Use Disorders: A Systematic Review and Meta-Analysis. J Dual Diagn. 2019 Sep;15(3):192–206.

29. Astuti Y, Wardhana A, Watkins J, Wulaningsih W, PILAR Research Network. Cigarette smoking and telomere length: A systematic review of 84 studies and meta-analysis. Environ Res. 2017;158:480–9.

30. Navarro-Mateu F, Rubio-Aparicio M, Cayuela P, Álvarez F-J, Roca-Vega A, Chirlaque MD, et al. The association of telomere length with substance use disorders: systematic review and meta-analysis protocol. Syst Rev. 2019 01;8(1):298.

31. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med. 2009 Jul 21;6(1549-1676 (Electronic)):e1000097.

32. Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. JAMA. 2000 Apr 19;283(0098-7484 (Print)):2008–12.

33. Wells G, Shea B, O'Connell D, Peterson J, Welch V, Losos M, et al. Ottawa Hospital Research Institute [Internet]. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. 2015 [cited 2020 May 10]. Available from: http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp

34. Jüni P, Witschi A, Bloch R, Egger M. The hazards of scoring the quality of clinical trials for meta-analysis. JAMA. 1999 Sep 15;282(11):1054–60.

35. Hedges L, Olkin I. Statistical Methods for Meta-Analysis. Orlando, USA: Academic Press; 1985.

36. Cohen J. Statistical Power Analysis for the Behavioral Sciences. New Yersey: Lawrence Erlbaum Associates; 1988.

37. Hartung J, Knapp G. On tests of the overall treatment effect in meta-analysis with normally distributed responses. Stat Med. 2001 Jun 30;20(12):1771–82.

38. Rubio-Aparicio M, López-López JA, Sánchez-Meca J, Marín-Martínez F, Viechtbauer W, Van den Noortgate W. Estimation of an overall standardized mean difference in random-effects meta-analysis if the distribution of random effects departs from normal. Res Synth Methods. 2018 Sep;9(3):489–503.

39. Sanchez-Meca J, Marin-Martinez F. Confidence intervals for the overall effect size in random-effects meta-analysis. PsycholMethods. 2008 Mar;13(1082-989X (Print)):31–48.

40. IntHout J, Ioannidis JPA, Rovers MM, Goeman JJ. Plea for routinely presenting prediction intervals in meta-analysis. BMJ Open. 2016 12;6(7):e010247.

41. Viechtbauer W. Confidence intervals for the amount of heterogeneity in meta-analysis. Stat Med. 2007 Jan 15;26(1):37–52.

42. Mathur MB, Van der Weele TJ. New metrics for meta-analyses of heterogeneous effects. Stat Med. 2019 Apr 15;38(8):1336–42.

43. Rubio-Aparicio M, López-López JA, Viechtbauer W, Marín-Martínez F, Botella J, Sánchez-Meca J. Testing Categorical Moderators in Mixed-Effects Meta-analysis in the Presence of Heteroscedasticity. J Exp Educ. 2020 Feb 3;88(2):288–310.

44. Knapp G, Hartung J. Improved tests for a random effects meta-regression with a single covariate. Stat Med. 2003 Sep 15;22(17):2693–710.

45. López-López JA, Marín-Martínez F, Sánchez-Meca J, Van den Noortgate W, Viechtbauer W. Estimation of the predictive power of the model in mixed-effects meta-regression: A simulation study. Br J Math Stat Psychol. 2014 Feb;67(1):30–48.

46. Duval S, Tweedie R. Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. Biometrics. 2000 Jun;56(2):455–63.

47. Rothstein HR, Sutton AJ, Borenstein M. Publication bias in meta-analysis: Prevention, assessment, and adjustments. Chichester, UK: Wiley.; 2005.

48. Stanley TD, Doucouliagos H. Meta-regression approximations to reduce publication selection bias. Res Synth Methods. 2014 Mar;5(1):60–78.

49. Viechtbauer W. Conducting Meta-Analyses in R with the metafor Package. J Stat Softw. 2010 Aug 5;36(1):1–48.

50. Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, et al. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. J Clin Epidemiol. 2011 Apr;64(4):383–94.

51. Tannous J, Mwangi B, Hasan KM, Narayana PA, Steinberg JL, Walss-Bass C, et al. Measures of possible allostatic load in comorbid cocaine and alcohol use disorder: Brain white matter integrity, telomere length, and anti-saccade performance. PLoS ONE. 2019;14(1):e0199729.

52. Martins de Carvalho L, Wiers CE, Manza P, Sun H, Schwandt M, Wang G-J, et al. Effect of alcohol use disorder on cellular aging. Psychopharmacology (Berl). 2019 Nov;236(11):3245–55.

53. Dixit S, Whooley MA, Vittinghoff E, Roberts JD, Heckbert SR, Fitzpatrick AL, et al. Alcohol consumption and leukocyte telomere length. Sci Rep. 2019 05;9(1):1404.

54. Aida J, Yokoyama A, Izumiyama N, Nakamura K, Ishikawa N, Poon SS, et al. Alcoholics show reduced telomere length in the oesophagus. J Pathol. 2011 Feb;223(3):410–6.

55. Yamaki N, Matsushita S, Hara S, Yokoyama A, Hishimoto A, Higuchi S. Telomere shortening in alcohol dependence: Roles of alcohol and acetaldehyde. J Psychiatr Res. 2019;109:27–32.

56. Mohamed MA, Ibrahim KS, Mahdy-Abdallah H, Mohamed HA. Leukocyte telomere length in heavy tobacco and marijuana Egyptian smokers. Int J ChemTech Res. 2016;9(4):501–8.

57. Pavanello S, Hoxha M, Dioni L, Bertazzi PA, Snenghi R, Nalesso A, et al. Shortened telomeres in individuals with abuse in alcohol consumption. Int J Cancer. 2011 Aug 15;129(4):983–92.

58. Levandowski ML, Tractenberg SG, de Azeredo LA, De Nardi T, Rovaris DL, Bau CHD, et al. Crack cocaine addiction, early life stress and accelerated cellular aging among women. Prog Neuropsychopharmacol Biol Psychiatry. 2016 Nov 3;71:83–9.

59. Yang Z, Ye J, Li C, Zhou D, Shen Q, Wu J, et al. Drug addiction is associated with leukocyte telomere length. Sci Rep. 2013;3:1542.

60. Savolainen K, Raikkonen K, Kananen L, Kajantie E, Hovatta I, Lahti M, et al. History of mental disorders and leukocyte telomere length in late adulthood: the Helsinki Birth Cohort Study (HBCS). J Psychiatr Res. 2012;46(10):1346–53.

61. Ryan R, Hill S. How to GRADE the quality of the evidence. Cochrane Consumers and Communication Group. Version 3.0). [Internet]. 2016 [cited 2020 May 14]. Available from: http://cccrg.cochrane.org/author-resources.

62. Malouff JM, Schutte NS. A meta-analysis of the relationship between anxiety and telomere length. Anxiety Stress Coping. 2017 May;30(3):264–72.

63. Mallen C, Peat G, Croft P. Quality assessment of observational studies is not commonplace in systematic reviews. J Clin Epidemiol. 2006 Aug;59(8):765–9.

64. Little J, Higgins JP, Ioannidis JP, Moher D, Gagnon F, von EE, et al. STrengthening the REporting of Genetic Association studies (STREGA): an extension of the STROBE Statement. AnnInternMed. 2009 Feb 3;150(1539-3704 (Electronic)):206–15.

65. von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. PLoSMed. 2007 Oct 16;4(1549-1676 (Electronic)):e296.

66. Li Z, He Y, Wang D, Tang J, Chen X. Association between childhood trauma and accelerated telomere erosion in adulthood: A meta-analytic study. J Psychiatr Res. 2017 Oct;93:64–71.

67. Oliveira BS, Zunzunegui MV, Quinlan J, Fahmi H, Tu MT, Guerra RO. Systematic review of the association between chronic social stress and telomere length: A life course perspective. Ageing Res Rev. 2016 Mar;26:37–52.

68. Pepper GV, Bateson M, Nettle D. Telomeres as integrative markers of exposure to stress and adversity: a systematic review and meta-analysis. R Soc Open Sci. 2018 Aug;5(8):180744.

69. Lai T-P, Wright WE, Shay JW. Comparison of telomere length measurement methods. Philos Trans R Soc Lond B Biol Sci. 2018 05;373(1741).

70. Lin J, Smith DL, Esteves K, Drury S. Telomere length measurement by qPCR - Summary of critical factors and recommendations for assay design. Psychoneuroendocrinology. 2019 Jan;99:271–8.

71. Savolainen K, Räikkönen K, Kananen L, Kajantie E, Hovatta I, Lahti M, et al. History of mental disorders and leukocyte telomere length in late adulthood: the Helsinki Birth Cohort Study (HBCS). J Psychiatr Res. 2012 Oct;46(10):1346–53.

72. Tomita K. How long does telomerase extend telomeres? Regulation of telomerase release and telomere length homeostasis. Curr Genet. 2018 Dec;64(6):1177–81.

73. Dugdale HL, Richardson DS. Heritability of telomere variation: it is all about the environment! Philos Trans R Soc Lond B Biol Sci. 2018 05;373(1741).

74. Bateson M, Nettle D. Why are there associations between telomere length and behaviour? Philos Trans R Soc Lond B Biol Sci. 2018 05;373(1741).

75. Rasgon N, Lin KW, Lin J, Epel E, Blackburn E. Telomere length as a predictor of response to Pioglitazone in patients with unremitted depression: a preliminary study. Transl Psychiatry. 2016 Jan 5;6:e709.

76. Hough CM, Bersani FS, Mellon SH, Epel ES, Reus VI, Lindqvist D, et al. Leukocyte telomere length predicts SSRI response in major depressive disorder: A preliminary report. Mol Neuropsychiatry. 2016 Jul;2(2):88–96.

77. Squassina A, Pisanu C, Congiu D, Caria P, Frau D, Niola P, et al. Leukocyte telomere length positively correlates with duration of lithium treatment in bipolar disorder patients. Eur Neuropsychopharmacol J Eur Coll Neuropsychopharmacol. 2016;26(7):1241–7.

78. Schutte NS, Malouff JM. A meta-analytic review of the effects of mindfulness meditation on telomerase activity. Psychoneuroendocrinology. 2014 Apr;42:45–8.

79. Arsenis NC, You T, Ogawa EF, Tinsley GM, Zuo L. Physical activity and telomere length: Impact of aging and potential mechanisms of action. Oncotarget. 2017 Jul 4;8(27):45008–19.

Table 1: Main characteristics of included studies

First author, year	Country	Study design	Cases		Controls		Age		Substance	SUD (†) instrument	Overall association
			N	Males N (%)	N	Males N (%)	Cases Mean (SD/range)	Control Mean (SD/range)			
Aida, 2011	Japan	Case-Control	26	26	24	12	61.2 (44-82)	73.3 (41-95)	Alcohol	DSM-IV	Shorter TL in the oesophagial epithelium of cases.
Pavanello, 2011	Italy	Case-Control	200	200	257	257	38 35-75)	44 (25-62)	Alcohol & smoking	DSM-IV-TR	Shorter TL in alcohol abusers compared to controls
Savolainen, 2012	Finland	Nested Case-Control (‡)	40	0	1840	842	-	61.5 (2.9)	Not described	Finnish Hospital Discharge Register (ICD-9 and 10 and DSM-III-R classification systems)	Participants hospitalized for any mental or substance use disorders had longer TL than non-hospitalized controls.
Yang, 2013	China	Case-Control	415	199	508	210	33.79 (7.6))	34.46 (8.16)	Heroin, Morphine, Methadone, Tramadol, Marijuana, MDMA, Methamphetamine, Ketamine, Triazolam	Standarized questionnaires and protocols not described	Drug abusers exhibited significantly shorter TLs than controls.
Mohamed 2016	Egypt	Case-Control	30	30	NS(ϕ)= 30 S(ϕ)= 30	NS(ϕ)= 30 S(ϕ)= 30	41.83 (6.94)	NS(ϕ)=40.53 (7.52) S(ϕ)=41.07 (6.98)	Tabacco & marihuana	Not described (#)	TL was shorter in marijuana smokers group than in smokers or non-smokers groups.
Levandowski, 2016	Brasil	Case-Control	34	0	49	0	26.2 (6.3)	68.3 (7.4)	Cocaine	Clinical interview and semi-structured clinical interview following DMS-IV criteria, CSSA, ASI-6	TL is shorter in Crack with/without early life stress than in controls.
Tannous, 2019	USA	Case-Control	24	18	25	17	46.96 (7.66)	43.76 (6.62)	Alcohol & Cocaine	SCID-IV, ASI, KMSK	No significant TL differences in comorbid cocaine and alcohol use disorder.
Yamaki, 2019	Japan	Case-Control	134	134	121	121	58.7 (9.7)	59 (10.2)	Alcohol	Kurihama-Alcoholism Screening Test & DSM IV criteria	TL was almost 50% shorter in patients with alcohol dependence (AD) compared to controls.
Martins, 2019	USA	Case-Control	260	187	449	248	44.06 (0.73)	33.32 (0.56)	Alcohol	SCID, AUDIT, ADS	TL is shorter in participants with alcohol use disorders compared to healthy controls.
Dixit, 2019 (&) (A)	USA	Nested Case-Control (‡)	627	527	321	246	67.1 (10.8)	66.1 (11.2)	Alcohol	AUDIT-C	No association between TL and Alcohol consumers
(B)	USA	Nested Case-Control	790	376	883	314	74.5 (5)	75.1 (5.5)	Alcohol	Self reported	No association between TL and Alcohol consumers

(†) Note: SUD: Substance Use Disorder; AUDIT: Alcohol Use Disorders Identification Test; ICD: International Statistical Classification of Diseases and Related Health Problems; DSM: Diagnostic and Statistical Manual of Mental Disorders; CSSA: Cocaine Selective Severity Assessment; ASI-6: Addiction Severity Index version 6; SCID: Structured Clinical Interview for DSM-IV; ASI: Addiction Severity Index; KMSK: Kreek-McHugh-Schluger-Kellogg scale; ADS: Alcohol Dependence Scale; TL: Telomere length

(&) Dixit et al (2019) describes results of two other independent studies: (A) The Heart and Soul Study and (B) The Cardiovascular Health Study (B).

(‡) Only 5-year follow-up measurement of telomere length in cohorts was included in the analyses.

(#) Most of participants were recruited from a hospital. It was confirmed by measurement of delta-9-tetrahydrocannabinol (THC) in urine samples.

(ϕ) NS: Non-smokers; S: Smokers.

Table 2: Telomere measurement description in included studies

First Autor, year		Cases Telomere length (T/S ratio) Mean	Cases Telomere length (T/S ratio) SD	Controls Telomere length (T/S ratio) Mean	Controls Telomere length (T/S ratio) SD	Original TL measurement	Technique used	Tissue source	Blind assay assessment reported	Genetic quality control	Genetic quality control description
Aida, 2011		1,22	0,6	1,64	0,6	Normalized telomere-to-centromere ratio (NTCR)	Q-FISH	Oesophageal mucosa (basal and parabasal cells)	NR	**Yes**	Control for variations in sample preparation.
Pavanello, 2011		0,45	0.12	1.06	0.82	T/S ratio	Multiplex real-time qPCR	Leukocytes	**Yes**	**Yes**	All samples run in triplicate and the average of the three T/S ratio measurements was used. Repetition of the assay for 20 samples in two different ways.
Savolainen, 2012		NR	NR	NR	NR	T/S ratio	Real-time qPCR	Leukocytes	NR	**Yes**	Triplicates with amplification curve standard deviations above 0,5 at the threshold level were omitted.
Yang, 2013		0,78	0,18	0,84	0,21	T/S ratio	qPCR	Leukocytes	NR	**Yes**	All samples were run in duplicate and evaluated correlation. Samples with a CV>2% were excluded and re-run. To test the reproducibility of the assay, multivariate samples were randomly chosen and run again.
Mohamed 2016		0,61	0,09	0,77	0,11	T/S ratio	qPCR	Leukocytes	NR	**Yes**	In duplicate 10 ml reaction within the same plate and calibrator genomic DNA in each run of the samples.
Levandowski, 2016	(A)	1,33	0,16	1,5	0,42	T/S ratio	qPCR	Peripheral blood	NR	**Yes**	Prior to the experiment, primer sets were tested thoroughly to determine reaction efficiency, specificity, and the absence of primer-dimers.
	(B)	1,19	0,21	1,5	0,42	T/S ratio	qPCR	Peripheral blood	"	**"**	
Tannous, 2019		0,93	0.34	1,13	0.70	T/S ratio	PCR	Leukocytes	NR	**Yes**	DNA samples run in duplicate
Yamaki, 2019		1,49 (†)	0.626	3,96(†)	1.778	kbp (†)	Telo TAGGG Telomere Length Assay Kit	Leukocytes	NR	NR	-
Martins, 2019		1,06	0,00	1,14	0,01	T/S ratio	Monochrome multiplex qPCR	Whole blood	NR	**Yes**	Reactions were pipetted in triplicate for the standard curves and in duplicate for the other samples. In all reactions, a negative control without cDNA template (NTC) was tested.
Dixit, 2019	(A)	0,83 (†)	0.149	0,84(†)	0.145	kpb (†)	qPCR	Leukocytes	NR	**Yes**	T/S ratio measured in duplicate and the averaged was used for each participant.
	(B)	1,26 (†)	0.249	1,27(†)	0.261	kbp (†)	Southern blot analysis of terminal restriction fragment lengths	Leukocytez	**Yes**	**Yes**	Telomere measurements were performed in duplicate

Note: kbp; kilobase pairs; qPCR: quantitative Polymerasa Chain Reaction; Q-FISH: Quantitative Fluorescence In Situ Hybridazation; NR, not reported
(†) Base pairs (bp) were transformed to T/S ratio using the formula: bp = 3274 + 2413*(T/S) (Dixit et al, 2019)

Table 3: Quality characteristics of included studies

First Autor, year	NEW OTAWA-CASTLE SCALE for case-control studies									Evaluate …				Control for … (Matched or Controlled)*			
	SELECTION				COMPARABILITY	EXPOSURE			NOS score								
	1) Case definition	2) Representativeness of cases	3) Selection of controls	4) Definition of controls	1) Comparability	1) Ascertainment of Exposure	2) Same method of ascertainment	3) Non-Response rate		Psychiatric comorbidity (Yes/No)	Physical comorbidity (Yes/No)	Childhood adversities (Yes/No)	Other stressful event exposure (Yes/No)	Age	Sex	Smoking	Other covariates controlled
Aida, 2011	✸	-	-	-	-	-	-	-	1/9	No	No	No	No	No	No	No	-
Pavanello, 2011	✸	-	-	✸	✸✸	-	✸	-	5/9	No	No	No	**Yes**	**Yes (C)**	**Yes (M)**	**Yes (C)**	BMI, vegetable intake, and jobs with elevate risk of accident.
Savolainen, 2012	✸	✸	✸	✸	✸✸	-	✸	-	7/9	**Yes**	**Yes**	No	No	**Yes (C)**	**Yes (C)**	**Yes (C)**	Diabetes mellitus, BMI, alcohol consumption and coronary heart disease.
Yang, 2013	-	-	-	✸	✸✸	-	-	-	3/9	**Yes**	**Yes**	No	**Yes**	**Yes (C)**	**Yes (C)**	**Yes (M)**	Participant self-assessment indicated that all of them were free of serious illness (infectious disease, cardiovascular diseases, mental disorders and cancer).
Mohamed 2016	-	-	-	✸	✸✸	✸	✸	-	5/9	**Yes**	**Yes**	No	No	**Yes (M)**	**Yes (M)**	**Yes (M)**	Socioeconomic status (M).
Levandowski, 2016	✸	-	-	✸	✸✸	-	✸	-	5/9	**Yes**	**Yes**	**Yes**	No	**Yes (C)**	**Yes (M)**	No	BMI and education level.
Tannous, 2019	✸	-	✸	-	✸✸	-	✸	-	5/9	No	No	No	No	**Yes (C)**	No	No	Education.
Yamaki, 2019	-	-	✸	✸	✸✸	-	-	-	5/9	No	**Yes**	No	No	**Yes (C)**	No	**Yes (C)**	Alcohol consumption, and cancer.
Martins, 2019	✸	-	✸	✸	✸✸	-	✸	-	6/9	**Yes**	No	**Yes**	**Yes**	**Yes (C)**	**Yes (C)**	Yes (C)	BMI, years of education, and African ancestry,.
Dixit, 2019 (A)	✸	-	✸	✸	✸✸	-	✸	-	6/9	No	**Yes**	No	No	**Yes (C)**	**Yes (C)**	**Yes (C)**	Race, BMI, waist-hip ratio, number of pack years & a group of medical conditions (†) and inflammatory markers and omega-3 fatty acid levels.
(B)	-	-	✸	✸	✸✸	-	✸	-	5/9	No	**Yes**	No	No	**Yes (C)**	**Yes (C)**	**Yes (C)**	

Note: BDI: Beck Depression Inventory; MHI: Mental Health Index; VS: Vitality Scale; BMI: body mass index; CTQ: Childhood Trauma Questionnaire; ELS: Early Life Stress Questionnaire; NR, not reported
(†) Medical conditions: diabetes, hypertension, coronary artery disease, prior myocardial infarction, heart failure, prior stroke and liver disease. (*) The statistical analyses applied in each study are described in Supplementary Table S2. (M) = control by matching; (C) = Statistical control.

Table 4. Results of the subgroup analyses for the Newcastle-Ottawa Scale (NOS)[a] and methodological characteristics on the effect sizes.

Moderator variable	*N*	*k*	d_+	95% CI d_L	95% CI d_U	ANOVA results
NOS-S1) Case definition						
No	4,814	5	-0.66	-1.32	0.004	$F(1, 10) = 0.02$, $p = .896$
Yes	2,389	7	-0.61	-1.18	-0.03	$R^2 = 0$ $Q_W(10) = 246.71$, $p < .001$
NOS-S2) Representativeness of cases						
No	5,323	11	-0.72	-1.12	-0.31	$F(1, 10) = 2.84$, $p = .123$
Yes	1,880	1	0.32	-0.99	1.64	$R^2 = .14$ $Q_W(10) = 239.53$, $p < .001$
NOS-S3) Selection of controls						
No	1,993	8	-0.94	-1.34	-0.53	$F(1, 10) = 8.31$, ***p* = .016**
Yes	5,210	4	-0.07	-0.60	0.46	$R^2 = .43$ $Q_W(10) = 152.54$, $p < .001$
NOS-S4) Definition of controls						
No	50	1	-0.69	-2.27	0.90	$F(1, 10) = 0.01$, $p = .933$
Yes	7,153	11	-0.62	-1.08	-0.17	$R^2 = 0$ $Q_W(10) = 254.98$, $p < .001$
NOS-C1) Comparability						
No	50	1	-0.69	-2.27	0.90	$F(1, 10) = 0.01$ $p = .933$
Yes	7,153	11	-0.62	-1.08	-0.17	$R^2 = 0$ $Q_W(10) = 254.98$, $p < .001$
NOS-E1) Ascertainment of exposure						
No	7,113	11	-0.56	-0.97	-0.14	$F(1, 10) = 1.91$, $p = .196$
Yes	90	1	-1.48	-2.91	-0.05	$R^2 = .08$ $Q_W(10) = 234.82$, $p < .001$
NOS-E2) Same method of ascertainment						
No	1,221	3	-0.96	-1.79	-0.14	$F(1, 10) = 1.09$, $p = .321$
Yes	5,982	9	-0.52	-0.99	-0.04	$R^2 = .01$ $Q_W(10) = 233.99$, $p < .001$
Controls without SUD?[b]						
No	2,729	4	-0.60	-1.36	0.15	$F(1,10) = 0.01$, $p = .925$
Yes	4,474	8	-0.64	-1.17	-0.11	$R^2 = 0$ $Q_W(10) = 228.51$, $p < .001$
Blinded assessors?						
Not reported	5,073	10	-0.65	-1.13	-0.18	$F(1, 10) = 0.08$, $p = .785$
Yes	2,130	2	-0.51	-1.54	0.51	$R^2 = 0$ $Q_W(10) = 249.02$, $p < .001$
Genotyping quality control[c]						
No	255	1	-1.88	-3.04	-0.73	$F(1, 10) = 6.54$, ***p* = .028**
Yes	6,948	11	-0.50	-0.86	-0.15	$R^2 = .37$ $Q_W(10) = 145.34$, $p < .001$

N = total sample size. *k* = number of studies. d_+ = average effect size. d_L and d_U = lower and upper confidence limits for d_+. *F* = *F*-statistic for testing the significance of the moderator. R^2 = proportion of variance accounted for by the moderator. Q_W = statistic for testing the model misspecification.

[a] NOS-E3: No missing data or similar attrition for cases and controls. This last item was not analysed because no study fulfilled it.

[b] Controls were assessed for absence of SUD with a validated instrument.

[c] Reporting of quality control procedures in genotyping methods.

Table 5. Subgroup analyses for different characteristics on the effect sizes.

Moderator variables	*N*	*k*	d_+	95% CI d_L	95% CI d_U	ANOVA results
Substantive variables						
Type of substance:						
Alcohol	4,092	6	-0.68	-1.28	-0.08	$F(2, 8) = 0.25$, $p = .788$
Other	1,182	4	-0.86	-1.62	-0.10	$R^2 = 0$ $Q_W(8) = 235.87$, $p < .001$
Alcohol + Cocaine	49	1	-0.35	-1.93	1.22	
SUD measurement:						
Clinical interview	4,614	10	-0.73	-1.18	-0.27	$F(1, 10) = 1.30$, $p = .280$
Self-report	2,589	2	-0.18	-1.14	0.78	$R^2 = .02$ $Q_W(10) = 213.34$, $p < .001$
Telomere measurement:						
qPCR	5,225	9	-0.55	-1.04	-0.06	$F(1, 10) = 0.53$, $p = .482$
Other[a]	1,978	3	-0.87	-1.72	-0.02	$R^2 = 0$ $Q_W(10) = 250.87$, $p < .001$
Source tissue:						
Leukocytes	6,268	8	-0.59	-1.15	-0.03	$F(2, 9) = 0.05$, $p = .953$
Other blood samples	885	3	-0.73	-1.67	0.20	$R^2 = 0$ $Q_W(9) = 242.02$, $p < .001$
Other tissue[b]	50	1	-0.69	-2.37	0.99	
Psychiatric comorbidity?:						
Not reported	3,432	6	-0.66	-1.26	-0.07	$F(2, 9) = 0.97$, $p = .415$
Excluded	1,182	4	-0.86	-1.60	-0.12	$R^2 = 0$ $Q_W(9) = 250.86$, $p < .001$
Assessed but not excluded	2,589	2	-0.09	-1.11	0.91	
Physical comorbidity?:						
Not reported	1,215	3	-0.62	-1.51	0.26	$F(2, 9) = 0.43$, $p = .660$
Excluded	1,232	5	-0.83	-1.53	-0.13	$R^2 = 0$ $Q_W(9) = 194.84$, $p < .001$
Assessed but not excluded	4,756	4	-0.40	-1.16	0.35	
Child trauma						
Not reported	6,318	9	-0.60	-1.09	-0.10	$F(1, 10) = 0.09$, $p = .771$
Assessed but not excluded	885	3	-0.73	-1.60	0.14	$R^2 = 0$ $Q_W(10) = 243.98$, $p < .001$
Other stressful exposures?:						
Not reported	5,121	9	-0.64	-1.15	-0.13	$F(1, 10) = 0.01$, $p = .917$
Assessed but not excluded	2,082	3	-0.59	-1.44	0.25	$R^2 = 0$ $Q_W(10) = 229.07$, $p < .001$
Contextual variables						
Ethnicity:						
Caucasian	2,337	2	-0.34	-1.42	0.74	$F(2, 8) = 0.25$, $p = .788$
Asian	1,221	3	-0.96	-1.87	-0.06	$R^2 = 0$ $Q_W(8) = 235.87$, $p < .001$
Arabic	90	1	-0.23	-1.00	0.54	
Mixed	3,379	4	-1.48	-3.10	0.14	
Country:						
Brazil	176	2	-0.87	-1.74	-0.002	$F(6, 5) = 3.15$, $p = .114$
China	916	1	-0.31	-1.35	0.73	$R^2 = .61$ $Q_W(5) = 42.50$, $p < .001$
Egypt	90	1	-1.48	-2.70	-0.26	
Finland	1,880	1	0.32	-0.78	1.43	
Italy	457	1	-0.98	-2.04	0.08	
Japan	305	2	-1.38	-2.21	-0.54	
USA	3,379	4	-0.22	-0.77	0.37	
Continent:						
Africa	90	1	-1.48	-2.98	0.02	$F(4, 7) = 1.30$, $p = .357$
Asia	1,221	3	-0.96	-1.80	-0.13	$R^2 = .09$ $Q_W(7) = 168.64$, $p < .001$
Europe	2,337	2	-0.34	-1.34	0.66	
North America	3,379	4	-0.23	-0.94	0.48	
South America	176	2	-0.86	-1.93	0.20	
Funding?:						
Yes	7,113	11	-0.55	-0.97	-0.14	$F(1, 10) = 1.91$, $p = .196$
Unclear	90	1	-1.48	-2.91	-0.005	$R^2 = .08$ $Q_W(10) = 234.82$, $p < .001$

Note: *N* = total sample size. *k* = number of studies. d_+ = average effect size. d_L and d_U = lower and upper confidence limits for d_+. *F* = *F*-statistic for testing the significance of the moderator. R^2 = proportion of variance accounted for by the moderator. Q_W = statistic for testing the model misspecification.
[a] 'Q-FISH' (Quantitative Fluorescence In Situ Hybridization), 'TAGGG Telomere Length Assay Kit', and 'Southern blot analysis of terminal restriction fragment lengths'. [b] 'Oesophageal mucosa'.

Table 6. Results of the mixed-effects meta-regressions for continuous moderators on the effect sizes.

	k	Min.	Max.	Mean	b_j	*t*	*p*	Q_E	*p*	R^2
Substantive variables										
Year	12	2011	2019	2016	-0.014	-0.23	.821	237.68	< .001	0
Total mean age	12	34.2	74.8	50.5	0.013	1.02	.330	197.38	< .001	.004
Case mean age	11	26.2	74.5	47.4	0.010	0.81	.436	176.60	< .001	0
Control mean age	12	33.3	75.1	55.7	0.006	0.48	.643	213.38	< .001	0
Total SD of age	10	2.9	21.9	9.6	-0.040	-0.97	.358	189.80	< .001	.006
Case SD of age	9	0.7	10.8	6.9	-0.046	-0.56	.593	193.22	< .001	0
Control SD of age	10	0.6	11.2	6.7	-0.079	-1.12	.297	207.12	< .001	.01
Total percent male	11	0	100	61.5	-0.006	-1.06	.315	184.03	< .001	.03
Case percent male	11	0	100	66.1	-0.004	-0.84	.425	190.74	< .001	0
Control percent male	12	0	100	56.1	-0.007	-1.34	.210	186.57	< .001	.08
Methodological variables										
Total sample size	**12**	**49**	**1,880**	**600**	**0.0007**	**3.34**	**.007**	**105.28**	**< .001**	**.54**
Mean age difference[a]	11	-42.1	10.7	-7.7	0.005	0.45	.662	230.31	< .001	0
SD of age difference[b]	9	-1.1	1.04	-0.2	0.094	0.23	.822	186.30	< .001	0
Percent male difference[c]	11	0	50	9.0	0.012	0.91	.384	200.70	< .001	0
NOS Total score[d]	12	1	6	4.6	0.161	1.26	.237	230.42	< .001	.07

Note: *k* = Number of studies. Min. and Max. = Minimum and Maximum values of the moderator variable. b_j =Regression coefficient of the moderator. *t* = Statistic for testing the significance of the moderator. Q_E = statistic for testing the model misspecification. R^2 =Proportion of variance accounted for by the moderator. In boldface is highlighted the moderator that reached statistical significance.

[a] Mean age difference = Mean age of cases minus mean age of controls.

[b] SD of age difference = Age SD of cases minus age SD of controls.

[c] Percent male difference = Percent male of cases minus percent male of controls.

[d] Range of NOS total score: 0 – 9.

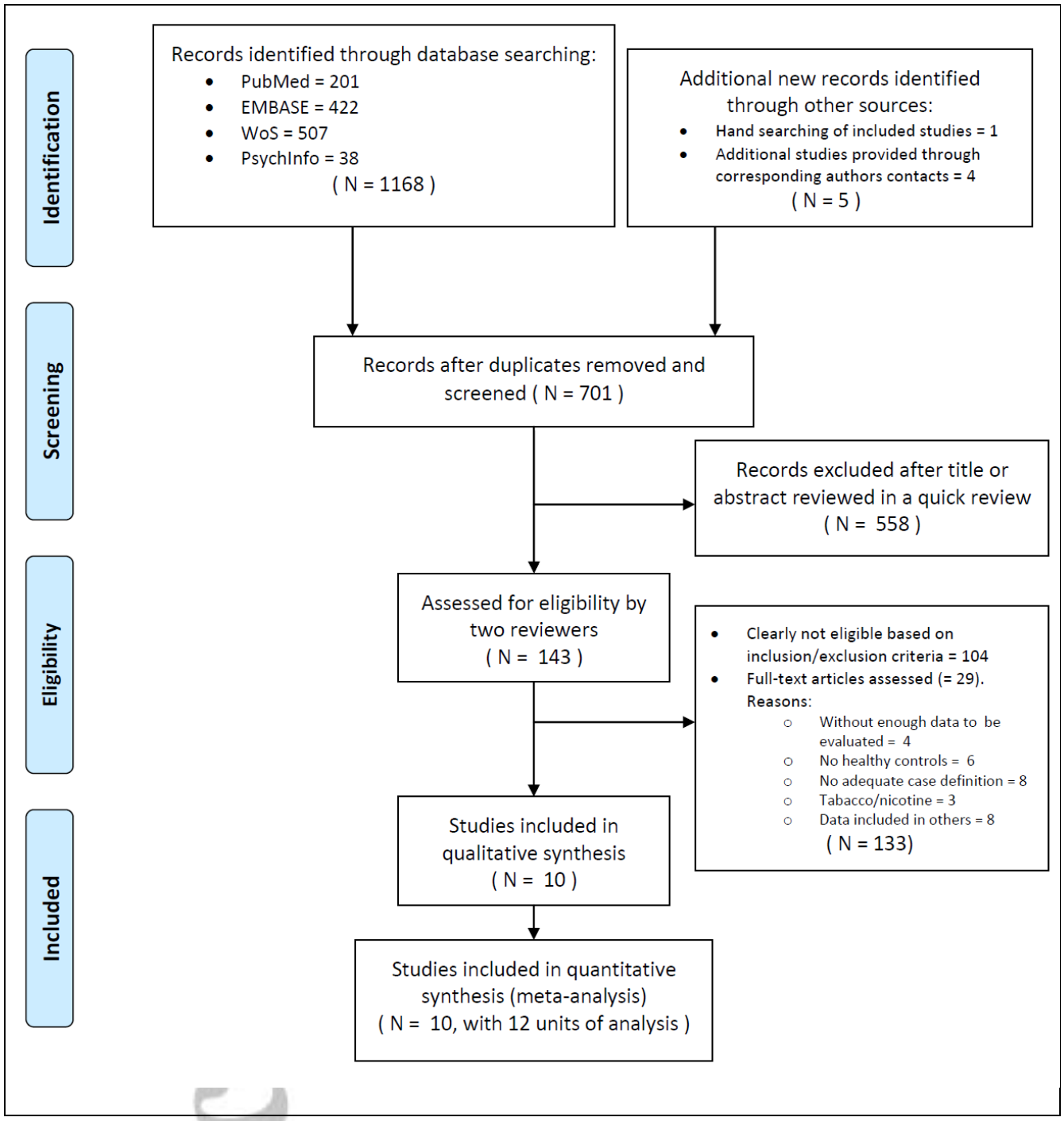


Figure 1. Flow chart of the Meta-Analysis of Telomere length and Substance Use Disorders. Adapted from Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(6): e1000097. doi:10.1371/journal.pone.0066227.g001

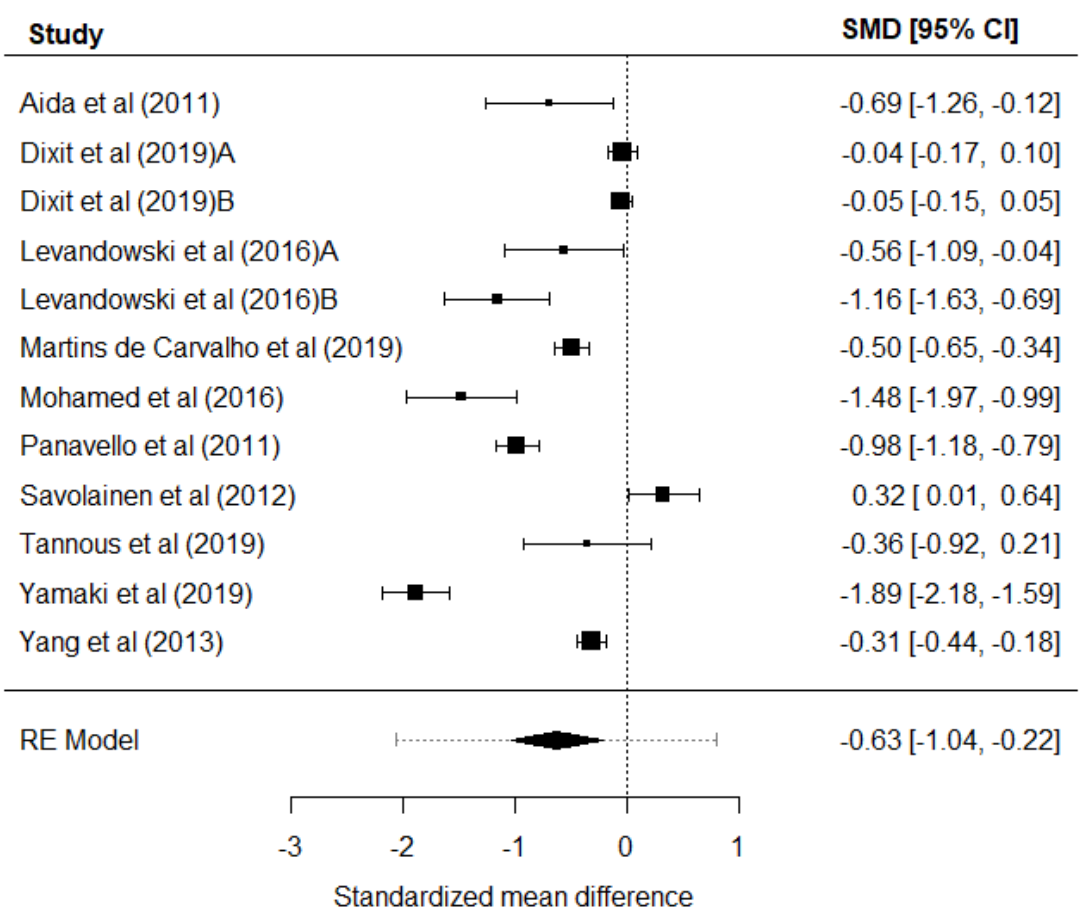


Figure 2. Forest plot of the standardized mean differences comparing average telomere length of SUD and control samples. SMD = Standardized Mean Difference.

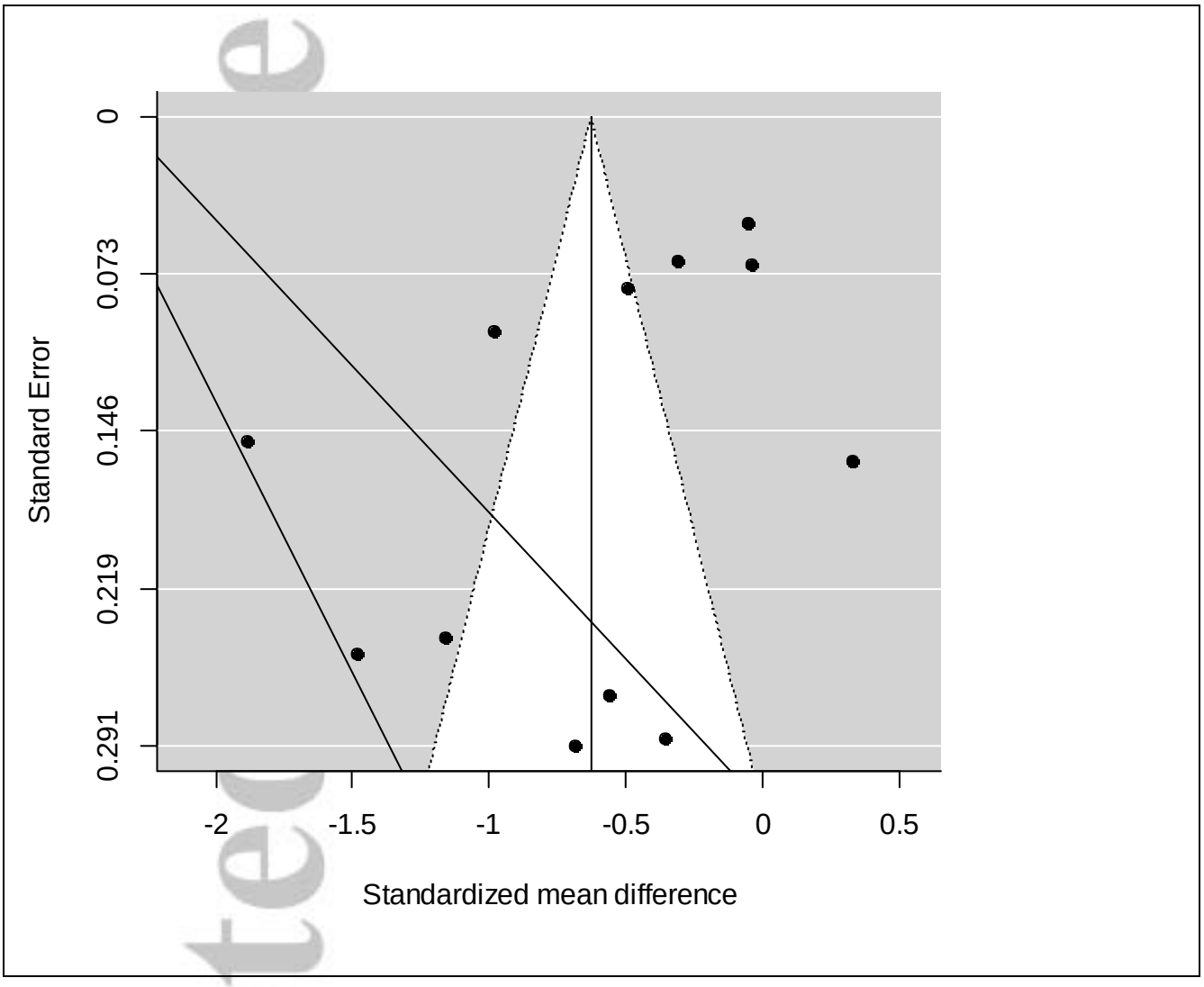
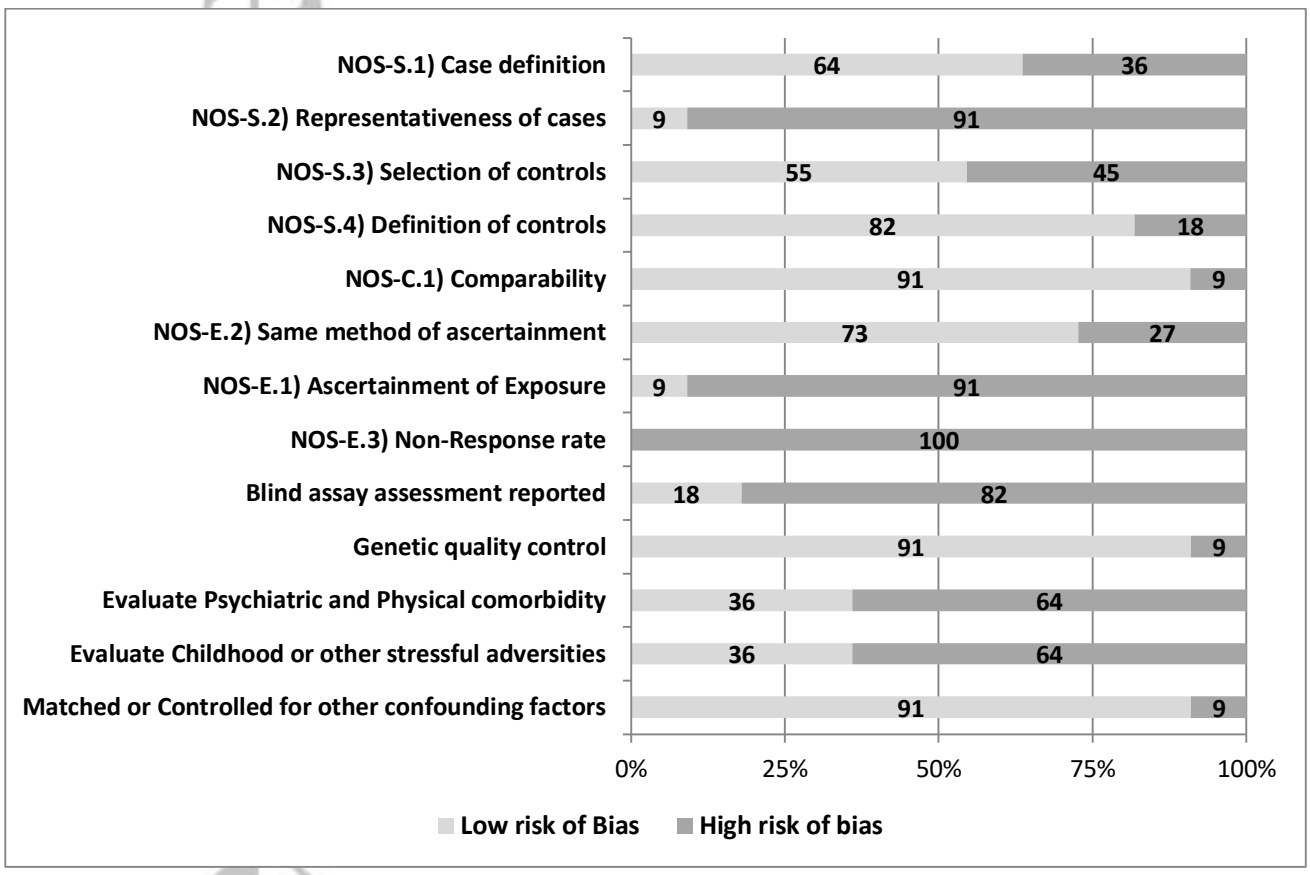


Figure 3. Funnel plot of the 12 standardized mean differences comparing average telomere length of SUD and control samples.



NOS: Newcastle-Otawa Scale for case-control studies. NOS-S: Selection; NOS-C: Comparability; and NOS-E: Exposure.

Figure 4: Risk of bias assessment of included studies.